

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/758,979A  
Source: 1Fw/6  
Date Processed by STIC: 7/12/06

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:50

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw

```

3 <110> APPLICANT: Ausubel, Frederick M.
4     Calderwood, Stephen B.
5     Garsin, Danielle A.
6     Mylonakis, Eleftherios
7     Sifri, Costi D.
9 <120> TITLE OF INVENTION: Enterococcal Virulence Factors
11 <130> FILE REFERENCE: 00786/408002
13 <140> CURRENT APPLICATION NUMBER: US 10/758,979A
14 <141> CURRENT FILING DATE: 2004-01-16
16 <150> PRIOR APPLICATION NUMBER: PCT/US02/22979
17 <151> PRIOR FILING DATE: 2002-07-18
19 <150> PRIOR APPLICATION NUMBER: US 60/306,212
20 <151> PRIOR FILING DATE: 2001-07-18
22 <160> NUMBER OF SEQ ID NOS: 47
24 <170> SOFTWARE: PatentIn version 3.3
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 2267
28 <212> TYPE: DNA
29 <213> ORGANISM: Enterococcus faecalis
31 <400> SEQUENCE: 1
32 gaaattgacc gtaaagtcaa tatttacaac attttaagcc tgaatgcgct aggtacgcta      60
34 tacgaactgg aaaaagatat gagaaaagcg caagtgtatt acgaaaaatc attacaagaa      120
36 ttggaacaat ttaaattaga atgttccttg gagcggttgta gaatttatta taattctgct      180
38 aaattctact cggaaatgaa agactaccaa aaaagtggtca ttttaagcga aaaagggatt      240
40 cagatttgct gtgacaaaca ctccatttat ttgctagatt atcttttata tgaaaaagcc      300
42 tttaacaaac aaatgctcgg ggaagacaca gccgatgact atcgccaagc ctattatttt      360
44 acacaatttt ttggcaatac ggaagtcctg caatatattg agaaagatat gaaagctttt      420
46 aatatttcct attaatTTaa tcaaaaagcc gataaaagct gaaaactcag tttttaccgg      480
48 ctttttgaaa aatataggca agttgctttt aaaaatcagc agtcacggtt acgataagca      540
50 agacgaagta tttaggagga tttaaaaatg aaaagagtaa tatgggttag acgtgattta      600
52 cgattacagg ataataaagc attagcacac gcgttacaaa attctgcagc tgatgaattg      660
54 attttattat tccaaatgaa tcctcaacaa tttattcaag aaagtgcata tcataacgct      720
56 ttttttgcaa gcttagcctc gttcaaagaa cgaatcgatc aagaggcaca tttacaaatc      780
58 atggtcggcg aaccattaga tttattttca cgtttgaaac gcaaaattacc cgattggcag      840
60 gccatttatt ttaatgaaga tacttggtggc tttggggcaa agcgggacca gcaagctatg      900
62 cgcttttttg aagaaaataa tattcagtct ttctcttttc aagatgccta tttgcatggc      960
64 tctgaagaaa ttaagaagaa cgatggcagc aagtaccaag tgtttacgcc ctattacaat      1020
66 aaatggaaag aggcgcctaa agaaacaccg attcctgttt cctatacagc tgaaaaaatt      1080
68 tttagtgcgt gtcctttttc agaagaggaa gcagcttatc gtgaacagat tgcgaggatt      1140
70 cctttaacac actatagtgt cggcgaagaa acagccagaa ggcgcttaaa tacttttatt      1200
72 gatcaaaaac ttcaatccta tgaaaaataag cgtgattttc cttatcagga tcaaacgagt      1260
74 catctgtcta cttttttaag aacgggagaa ctttcgattc gcaccatttg gcaagagctt      1320
76 gcactctgtc cttctagctt aagtaaagaa accttcaaaa aagaattagc ttggcgcgac      1380

```

## RAW SEQUENCE LISTING

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:50

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw

```

78 ttttacaata tgatctatag tgcgtttcca caacaaaaag aggaagctat tcaagaaaaa 1440
80 tttcgttata ttcaatggac aaatgaccca gaaatgtttg tcaagtggca aaaaggggag 1500
82 acggggtacc ctataattga tgccgcaatg cgacaactga atcaaactgg ttggatgcac 1560
84 aatcgcttaa gaatgattac tgcctctttt ttagttaaaa atttacacat cgattggcgt 1620
86 tggggtgaaa aatactttca aaaaatgttg attgactatg atgctgcca taatatcgg 1680
88 ggctggcaat gggctgcttc aacaggaacg gacgctgtcc cttattttcg gattttta 1740
90 ccaattatcc agtcaaaaaa atttgataat gacggccagt tcatcaaaaa atatgttcca 1800
92 gaacttaagc aagtgccaca aaagtatatt catcaaccaa atctaataa cgaagcctta 1860
94 caaacgcaat atcatgtaca tttaggagaa aattatccaa aacccattgt cgattatgca 1920
96 tcaagtaaaa aacaaacatt gtttctatat gaagcgagca aagaaattca tcaagaaatg 1980
98 aacaatccaa ggtttcaata aacagtaaac ccaactagct tagcaacaa cttgtaagg 2040
100 tagttgggca aattaattag tcgaaagaga agtgcaattt atcggctaaa gttttttatt 2100
102 ctattctagt taacataata tacattatac aaagtagagt aaaaagcatt gaaaagcaaa 2160
104 caaaaccagt ctttagttta tctagactgg ttttgtcacg tacgttatat aaattatgct 2220
106 tgttgcttga tggcttgtgc gacacgtgct ccatattctg gatttac 2267

```

109 &lt;210&gt; SEQ ID NO: 2

110 &lt;211&gt; LENGTH: 1434

111 &lt;212&gt; TYPE: DNA

112 &lt;213&gt; ORGANISM: Enterococcus faecalis

114 &lt;400&gt; SEQUENCE: 2

```

115 atgaaaagag taatatgggt tagacgtgat ttacgattac aggataataa agcattagca 60
117 cacgcgttac aaaattctgc agctgatgaa ttgattttat tattccaaat gaatcctcaa 120
119 caatttattc aagaaaagtgc taatcataac gctttttttg caagcttagc ctcgttcaaa 180
121 gaacgaatcg atcaagaggc acatttacaa atcatggtcg gcgaaccatt agatttattt 240
123 tcacgtttga aacgcaaatt acccgattgg caggccattt attttaatga agatacttgt 300
125 ggctttgggg caaagcggga ccagcaagct atgcgctttt ttgaagaaaa taatattcag 360
127 tctttctctt ttcaagatgc ctatttgcac ggctctgaag aaattaagaa gaacgatggc 420
129 agcaagtacc aagtgtttac gccctattac aataaatgga aagaggcgcc taaagaaaca 480
131 ccgattcctg tttcctatac agctgaaaaa atttttagtg cgtgtctttt tccagaagag 540
133 gaagcagctt atcgtgaaca gattgcgagg attcctttta cacactatag tgtcggcgaa 600
135 gaaacagcca gaaggcgctt aaatactttt attgatcaaa aacttcaatc ctatgaaaat 660
137 aagcgtgatt ttccttatca ggatcaaacg agtcatctgt ctactttttt aagaacggga 720
139 gaactttcga ttcgcaccat ttggcaagag cttgcatctg tgcttcttag cttaagtaaa 780
141 gaaaccttca aaaaagaatt agcttggcgc gacttttaca atatgatcta tagtgcgttt 840
143 ccacaacaaa aagaggaagc tattcaagaa aaatttcgtt atattcaatg gacaaatgac 900
145 ccagaaatgt ttgtcaagtg gcaaaaaggg gagacggggt accctataat tgatgccgca 960
147 atgcgacaac tgaatcaaac tggttggatg cacaatcgct taagaatgat tactgcctct 1020
149 tttttagtta aaaatttaca catcgattgg cgttgggggtg aaaaatactt tcaaaaaatg 1080
151 ttgattgact atgatgctgc caataatatc ggtggctggc aatgggctgc ttcaacagga 1140
153 acggacgctg tcccttattt tcggattttt aatccaatta tccagtcaaa aaaatttgat 1200
155 aatgacggcc agttcatcaa aaaatatgtt ccagaactta agcaagtgcc acaaaagtat 1260
157 attcatcaac caaatcta atgaacgaagc ttacaaacgc aatatcatgt acatttagga 1320
159 gaaaattatc caaaacccat tgtcgattat gcatcaagta aaaaacaaac attgtttcta 1380
161 tatgaagcga gcaagaaat tcatacaagaa atgaacaatc caaggtttca ataa 1434

```

164 &lt;210&gt; SEQ ID NO: 3

165 &lt;211&gt; LENGTH: 477

166 &lt;212&gt; TYPE: PRT

167 &lt;213&gt; ORGANISM: Enterococcus faecalis

169 &lt;400&gt; SEQUENCE: 3

## RAW SEQUENCE LISTING

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:50

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw

```

171 Met Lys Arg Val Ile Trp Phe Arg Arg Asp Leu Arg Leu Gln Asp Asn
172 1 5 10 15
175 Lys Ala Leu Ala His Ala Leu Gln Asn Ser Ala Ala Asp Glu Leu Ile
176 20 25 30
179 Leu Leu Phe Gln Met Asn Pro Gln Gln Phe Ile Gln Glu Ser Ala Asn
180 35 40 45
183 His Asn Ala Phe Phe Ala Ser Leu Ala Ser Phe Lys Glu Arg Ile Asp
184 50 55 60
187 Gln Glu Ala His Leu Gln Ile Met Val Gly Glu Pro Leu Asp Leu Phe
188 65 70 75 80
191 Ser Arg Leu Lys Arg Lys Leu Pro Asp Trp Gln Ala Ile Tyr Phe Asn
192 85 90 95
195 Glu Asp Thr Cys Gly Phe Gly Ala Lys Arg Asp Gln Gln Ala Met Arg
196 100 105 110
199 Phe Phe Glu Glu Asn Asn Ile Gln Ser Phe Ser Phe Gln Asp Ala Tyr
200 115 120 125
203 Leu His Gly Ser Glu Glu Ile Lys Lys Asn Asp Gly Ser Lys Tyr Gln
204 130 135 140
207 Val Phe Thr Pro Tyr Tyr Asn Lys Trp Lys Glu Ala Pro Lys Glu Thr
208 145 150 155 160
211 Pro Ile Pro Val Ser Tyr Thr Ala Glu Lys Ile Phe Ser Ala Cys Leu
212 165 170 175
215 Phe Pro Glu Glu Glu Ala Ala Tyr Arg Glu Gln Ile Ala Arg Ile Pro
216 180 185 190
219 Leu Thr His Tyr Ser Val Gly Glu Glu Thr Ala Arg Arg Arg Leu Asn
220 195 200 205
223 Thr Phe Ile Asp Gln Lys Leu Gln Ser Tyr Glu Asn Lys Arg Asp Phe
224 210 215 220
227 Pro Tyr Gln Asp Gln Thr Ser His Leu Ser Thr Phe Leu Arg Thr Gly
228 225 230 235 240
231 Glu Leu Ser Ile Arg Thr Ile Trp Gln Glu Leu Ala Ser Val Pro Ser
232 245 250 255
235 Ser Leu Ser Lys Glu Thr Phe Lys Lys Glu Leu Ala Trp Arg Asp Phe
236 260 265 270
239 Tyr Asn Met Ile Tyr Ser Ala Phe Pro Gln Gln Lys Glu Glu Ala Ile
240 275 280 285
243 Gln Glu Lys Phe Arg Tyr Ile Gln Trp Thr Asn Asp Pro Glu Met Phe
244 290 295 300
247 Val Lys Trp Gln Lys Gly Glu Thr Gly Tyr Pro Ile Ile Asp Ala Ala
248 305 310 315 320
251 Met Arg Gln Leu Asn Gln Thr Gly Trp Met His Asn Arg Leu Arg Met
252 325 330 335
255 Ile Thr Ala Ser Phe Leu Val Lys Asn Leu His Ile Asp Trp Arg Trp
256 340 345 350
259 Gly Glu Lys Tyr Phe Gln Lys Met Leu Ile Asp Tyr Asp Ala Ala Asn
260 355 360 365
263 Asn Ile Gly Gly Trp Gln Trp Ala Ala Ser Thr Gly Thr Asp Ala Val
264 370 375 380
267 Pro Tyr Phe Arg Ile Phe Asn Pro Ile Ile Gln Ser Lys Lys Phe Asp

```

## RAW SEQUENCE LISTING

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:50

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw

```

268 385          390          395          400
271 Asn Asp Gly Gln Phe Ile Lys Lys Tyr Val Pro Glu Leu Lys Gln Val
272          405          410          415
275 Pro Gln Lys Tyr Ile His Gln Pro Asn Leu Met Asn Glu Ala Leu Gln
276          420          425          430
279 Thr Gln Tyr His Val His Leu Gly Glu Asn Tyr Pro Lys Pro Ile Val
280          435          440          445
283 Asp Tyr Ala Ser Ser Lys Lys Gln Thr Leu Phe Leu Tyr Glu Ala Ser
284          450          455          460
287 Lys Glu Ile His Gln Glu Met Asn Asn Pro Arg Phe Gln
288 465          470          475
291 <210> SEQ ID NO: 4
292 <211> LENGTH: 2543
293 <212> TYPE: DNA
294 <213> ORGANISM: Enterococcus faecalis
296 <400> SEQUENCE: 4
297 ccttttgaaa atttagatga aggatttgat gtctacgcca ctcaagcctt taatgcgcca      60
299 gatggtcgtg cacttgcggt cagttggatt gggttgccag aaatcactta cccaagtgat      120
301 gtggagggtt gggcaaatgg cttaagtctg gttaaagaac tcacaattca caacgggaaa      180
303 ctatttcaat atccagtttc tgaaacagaa atgcttcgtc aatccgctac tactttatca      240
305 aatggctgcc atttcttatt tactgcttct tttgaattag aagtggatat tcccaaaaat      300
307 gagattgctt ttattcggct tttagcgaac gaaacgggtt caaaaggact tttaattaca      360
309 attgatacga ttcattgtaa aataaccctt gatcgaacat ttgctggcca atcttttctg      420
311 gaaaagtatg gcacaattcg tgaaactaaa attaggaaaa ataagtcagt tcagttaact      480
313 atttttgttg attgctctgt tgcagaaatc tatgtaaata aagggtgaaaa aacgatgact      540
315 ggtcgcttct ttccagataa agcgcgaacag tatcttcatt tatccaagac ggcaaaaagt      600
317 tgtttttatg agctggaaaa tacgaataat taggaatgat ggtgaatttt gatggtggtt      660
319 aaattaacgg atgtagcaaa gcttgctggg gtgagcccga caacggttaag ccgctgatt      720
321 aataattatg gttatcttag tcaaaaaaca attgataaag ttcattcaagc gatggaagaa      780
323 ttaattatc aacctaattg attagccaga agcctccaag gaaaaagtac gcagctgatt      840
325 ggtttagtct tcccttctgt tagtcatcca ttttttggtg aattaattga aacactggaa      900
327 agaaagctct ttgttcaagg atataaagtg attttatgtg atagtgaaaa agatccagaa      960
329 aaagagcgcg cctattttacg aatgctcgct gcaaataaag tggacgggtg aatcactggt      1020
331 agccataact tagctattaa cgaatatgaa aatgtttcac tacctattgt ttcctttgac      1080
333 cgtttcttgg cacctggcat tccaattgtc tcttcgcaaa actttcaagg gggccaaaaa      1140
335 gccactgaag ccttatttgc aagtgggtga caaaagattg caattattac tggtgctaata      1200
337 aacacaggcg cacctagcga ttatcgattg gctggttata aacaaacaat ggaaaaatat      1260
339 ggcgagaaaa aaacgattct acaaattgat aatgggacct caacaacatt aaaaaatcta      1320
341 gaaatcgaac gtttgcttca aaataaaaact gtagacggca tcttttgtac agatgatttg      1380
343 acagcaatta cagttatgaa tattgctcaa aaattgaaga tatccattcc tgaagaatta      1440
345 aaagtaattg gttatgatgg gacaaaatta atcaaaagaa ttgccccaca actatcaacc      1500
347 attgtgcagc caatcgacga gatgtgtgac gttatgattg acttactgct tcgtagaatg      1560
349 aaggatcctg atgttgact tgaggaaaat tatcctattc cgattcagct atcattgtct      1620
351 gaatcctgtt aaaaaagaca ccttttcagg tgtctttttt tattcttaata tcctcatttg      1680
353 cctattttcc cttaattcca aagcacagag attaaaataat acaacaacat gatccctaca      1740
355 aatacaccca gcattagtaa aacataccac caatttttgg cgctataggt ttcatttttta      1800
357 tttttttcgg ttcgtttccc ggcagcccat ttccataaat cgtaattatt catctctctc      1860
359 accaccttta atgggtttat taacttaagc ctaacacgga atgagagaac ttgaaaaaaa      1920
361 aaggatcacc cgcttatgca atggatgatc ctttttttggg cactattctt ctgataaccc      1980

```

## RAW SEQUENCE LISTING

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:50

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw

```

363 attatttttta ataactgatt ggtaccaata aaaactatct tttttaatgc gtcgtaaadc 2040
365 ttttaactca tggctcttcac gattaacgta aataaaaccg tagcgttttt tgaatccttg 2100
367 atgggaactt aaaatatcca tgaccgacca agggcaatag ccaaacaact caacaccgtc 2160
369 agaaatagcc gcatgacaag cagcaatatg atcatgcaaa tacgcaattc gataatcatc 2220
371 atgaattttt ccgtcctctg ttaggtgacg tgggtgacct aagccgtttt ctgtaatgat 2280
373 taacggcaaa cggtattgac gataataatc atttaaaact aaacgtagac cagttggatc 2340
375 aatttgggca ccatacttag aggttttttaa atgttgattt ttttcgattt taaaataacc 2400
377 gtatagatca aaatcaatgt ccttttcttt cgtacccaat ggggtgttgct catctgtcgg 2460
379 taaatagcta gcgactaacg tccgataata attaagagcg ataaaaatctg gtttagcggc 2520
381 tttcaaaaatt gcttgatctt ctg 2543

```

384 &lt;210&gt; SEQ ID NO: 5

385 &lt;211&gt; LENGTH: 981

386 &lt;212&gt; TYPE: DNA

387 &lt;213&gt; ORGANISM: Enterococcus faecalis

389 &lt;400&gt; SEQUENCE: 5

```

390 atggtgggta aattaacgga ttagcaaaag cttgctgggg tgagcccgac aacggtaagc 60
392 cgctgatta ataattatgg ttatcttagt caaaaaacaa ttgataaagt tcatcaagcg 120
394 atggaagaat taaattatca acctaattga ttagccagaa gcctccaagg aaaaagtacg 180
396 cagctgattg gtttagtctt cccttctggt agtcatccat tttttggtga attaattgaa 240
398 aactggaaa gaaagctctt tgttcaagga tataaagtga tttatgtga tagtgaaaaa 300
400 gatccagaaa aagagcgcg cttattacga atgctcgctg caaataaagt ggacgggtgta 360
402 atcactggta gccataactt agctattaac gaatatgaaa atgtttcact acctattggt 420
404 tcctttgacc gtttcttggc acctggcatt ccaattgtct cttcgcaaaa ctttcaaggg 480
406 ggccaaaaag ccactgaagc cttatttgca agtggcgac aaaaagattg aattattact 540
408 ggtgctaata acacaggcgc acctagcgat tatcgattgg ctgggtataa acaacaatg 600
410 gaaaaatatg gcgcagaaaa aacgattcta caaattgata atgggacctc aacaacatta 660
412 aaaaatctag aaatcgaacg tttgcttcaa aataaaactg tagacggcat cttttgtaca 720
414 gatgatttga cagcaattac agttatgaat attgctcaaa aattgaagat atccattcct 780
416 gaagaattaa aagtaattgg ttatgatggg acaaaattaa tcaaaagaat tgccccacaa 840
418 ctatcaacca ttgtgcagcc aatcgacgag atgtgtgacg ttatgattga cttactgctt 900
420 cgtagaatga aggatcctga tgttgcaact gaggaaaatt atcctattcc gattcagcta 960
422 tcattgtctg aatcctgtta a 981

```

425 &lt;210&gt; SEQ ID NO: 6

426 &lt;211&gt; LENGTH: 326

427 &lt;212&gt; TYPE: PRT

428 &lt;213&gt; ORGANISM: Enterococcus faecalis

430 &lt;400&gt; SEQUENCE: 6

```

432 Met Val Val Lys Leu Thr Asp Val Ala Lys Leu Ala Gly Val Ser Pro
433 1 5 10 15
436 Thr Thr Val Ser Arg Val Ile Asn Asn Tyr Gly Tyr Leu Ser Gln Lys
437 20 25 30
440 Thr Ile Asp Lys Val His Gln Ala Met Glu Glu Leu Asn Tyr Gln Pro
441 35 40 45
444 Asn Gly Leu Ala Arg Ser Leu Gln Gly Lys Ser Thr Gln Leu Ile Gly
445 50 55 60
448 Leu Val Phe Pro Ser Val Ser His Pro Phe Phe Gly Glu Leu Ile Glu
449 65 70 75 80
452 Thr Leu Glu Arg Lys Leu Phe Val Gln Gly Tyr Lys Val Ile Leu Cys
453 85 90 95

```

**VERIFICATION SUMMARY**

DATE: 07/12/2006

PATENT APPLICATION: US/10/758,979A

TIME: 09:29:51

Input Set : A:\00786.408002.ST25.txt

Output Set: N:\CRF4\07122006\J758979A.raw